5/13
2/7/04

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	Maria Castellanos et al.	Examiner: Wayne P. Amsbury
Serial No.:	09/944,919	Group Art Unit: 2171
Filed:	August 31, 2001	Docket: 1007912-1
Title:	METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT	

RESPONSE TO NOTICE OF NON-COMPLIANT AMENDMENT

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

RECEIVED

JAN 3 0 2004

Technology Center 2100

Sir:

This Communication is in response to the "Notice of Non-Compliant Amendment" (see enclosed copy) mailed on January 9, 2004.

The Amendment and Response filed on December 2, 2003 was deemed non-compliant because the amendments to the claims were not underlined, and thus the amendment did not satisfy CFR 1.121. With this Communication we have submitted a completed set of claims in response to the Notice of Non-Compliant Amendment.

Applicant believes the Amendment and Response is now compliant and is in proper condition for examination. A fee is not required with this Response, however at any time during the pendency of this application, please charge any fees required or credit any overpayment to Deposit Account 500471.

Response to Notice of Non-Compliant Amendment
Applicant: Maria Castellanos et al.
Serial No.: 09/944,919
Filed: August 31, 2001
Docket No.: 10007912-1
Title: METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT

Any inquiry regarding this Communication should be directed to Steven E. Dicke at Telephone No. (612) 573-2002.

Respectfully submitted,

Maria Castellanos et al.,

By their attorneys,

DICKE, BILLIG & CZAJA, PLLC
Fifth Street Towers, Suite 2250
100 South Fifth Street
Minneapolis, MN 55402
Telephone: (612) 573-2002
Facsimile: (612) 573-2005
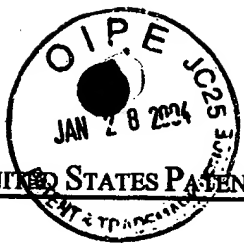
Dated: *January 26, 2004*
SED:jan

Steven E. Dicke
Steven E. Dicke
Reg. No. 38,431

CERTIFICATE UNDER 37 C.F.R. 1.8: The undersigned hereby certifies that this paper or papers, as described herein, are being deposited in the United States Postal Service, as first class mail, in an envelope address to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on this _____26_____ day of January, 2004.

By *Steven G. Dicke*
Name: Steven E. Dicke

2

UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/944,919	08/31/2001	Maria Castellanos	R-10007912 – /	3894

7590   01/09/2004

HEWLETT-PACKARD COMPANY
Intellectual Prperty Administration
P.O. Box 272400
Fort Collins, CO 80527-2400

RECEIVED

JAN 13 2004

HP LEGAL
IPA

EXAMINER
AMSBURY, WAYNE P

ART UNIT	PAPER NUMBER
2171	

DATE MAILED: 01/09/2004

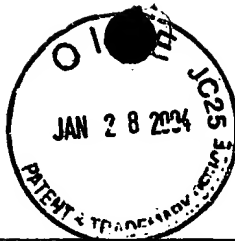Please find below and/or attached an Office communication concerning this application or proceeding.

US ACTION _____
OUR DATE _____
Paper Dated _____
OA _____ Final _____
Msg. Pt _____ Dwgs _____
Appeal _____ Issue Fee_____
Oth: Evaluate Doc
one month response!

RECEIVED

JAN 3 0 2004

Technology Center 2100

APPLICATION NO./ CONTROL NO.	FILING DATE	FIRST NAMED INVENTOR / PATENT IN REEXAMINATION	ATTORNEY DOCKET NO.

EXAMINER

ART UNIT	PAPER

4

DATE MAILED:

Please find below and/or attached an Office communication concerning this application or proceeding.

Commissioner for Patents

The proposed amendment of 12/5/03 is not in proper form. In particular, amendments to the claims are not underlined, and thus the amendment does not satisfy CFR 1.121. The timely submission under 37 CFR 1.129(a) filed on 12/5/03 is not fully responsive to the prior Office action because amendments to the claims are not underlined. Since the submission appears to be a *bona fide* attempt to provide a complete reply to the prior Office action, applicant is given a shortened statutory period of ONE MONTH or THIRTY DAYS from the mailing date of this letter, whichever is longer, to submit a complete reply. This shortened statutory period supersedes the time period set in the prior Office action. This time period may be extended pursuant to 37 CFR 1.136(a). If a notice of appeal and the fee set forth in 37 CFR 1.17(e) were filed prior to or with the payment of the fee set forth in 37 CFR 1.17(r), the payment of the fee set forth in 37 CFR 1.17(r) by applicant is construed as a request to dismiss the appeal and to continue prosecution under 37 CFR 1.129(a). The appeal stands dismissed.

RECEIVED

JAN 3 0 2004

Technology Center 2100

WAYNE AMSBURY
PRIMARY PATENT EXAMINER

WPA

PTO-90C (Rev.04-03)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	Maria Castellanos et al.	Examiner:	Wayne P. Amsbury
Serial No.:	09/944,919	Group Art Unit:	2171
Filed:	August 31, 2001	Docket No.:	10007912-1
Title:	METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT		

COMPLETE CLAIM SET ACCOMPANYING RESPONSE TO NOTICE OF NON-COMPLIANT AMENDMENT

IN THE CLAIMS

1. (Currently Amended) A computer-implemented method for mining a document containing dirty text comprising:

removing an instance of dirty text within said document to produce a cleaned document having a content; and

performing a data mining operation on said cleaned document thereby deriving relevant information from said cleaned document and providing a summary of the content of said document.

2. (Original) The method for mining a document containing dirty text as recited in Claim 1, wherein said removing further comprises replacing an instance of dirty text with a standard term.

3. (Original) The method for mining a document containing dirty text as recited in Claim 1, wherein said removing further comprises removing an instance of computer code from said document.

4. (Original) The method for mining a document containing dirty text as recited in Claim 1, wherein said removing further comprises removing a table from said document.

5. (Original) The method for mining a document containing dirty text as recited in Claim 1, wherein said performing a data mining operation further comprises identifying a sentence within said cleaned document by identifying a beginning and an end of said sentence.

COMPLETE CLAIM SET ACCOMPANYING RESPONSE TO NOTICE OF NON-COMPLIANT
AMENDMENT
Applicant: Maria Castellanos et al.
Serial No.: 09/944,919
Filed: August 31, 2001
Docket No.: H303.239.101
Title: METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT

6.   (Original) The method for mining a document containing dirty text as recited in Claim 5,
wherein said performing a data mining operation further comprises scoring and ranking said
sentence.

7.   (Original) The method for mining a document containing dirty text as recited in Claim 6,
wherein scoring said sentence further comprises:

        selecting scoring techniques operable for summarizing non-narrative, grammatically
incorrect text;

        selecting scoring techniques operable for summarizing narrative, grammatically correct
text; and

        using said scoring techniques to score said sentence.

8.   (Original) The method for mining a document containing dirty text as recited in Claim 7,
wherein said method further comprises generating a summary derived from said scored and
ranked sentences.

9.   (Original) The method for mining a document containing dirty text as recited in Claim 1,
wherein said method further comprises selecting a text mining component based upon said data
mining operation to be performed.

10.   (Original) The method for mining a document containing dirty text as recited in Claim 1,
wherein said method further comprises customizing said method by adjusting a parameter value.

11.   (Currently Amended) A computer system comprising:

        a bus;

        a memory unit coupled to said bus; and

        a processor coupled to said bus, said processor for executing a method for mining a
document containing dirty text comprising:

COMPLETE CLAIM SET ACCOMPANYING RESPONSE TO NOTICE OF NON-COMPLIANT
AMENDMENT
Applicant: Maria Castellanos et al.
Serial No.: 09/944,919
Filed: August 31, 2001
Docket No.: H303.239.101
Title: METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT

producing a cleaned document having a content comprising performing a general cleaning of said document by

_removing an instance of dirty text within said document including instances of misspelling and grammatical errors, and performing a domain and task specific cleaning of said document including removing instances of computer code and tables to produce a cleaned document; and

performing a data mining operation on said cleaned document including providing a summary of the content of said document.

12.   (Currently Amended) The computer system as recited in Claim 11, wherein said removing further comprises replacing an instance of dirty text with ~~an~~a standard term.

13.   (Cancelled)

14.   (Cancelled)

15.   (Original) The computer system as recited in Claim 11, wherein said performing a data mining operation further comprises identifying a sentence within said cleaned document by identifying a beginning and an end of said sentence.

16.   (Original) The computer system as recited in Claim 15, wherein said performing a data mining operation further comprises scoring and ranking said sentence.

17.   (Original) The computer system as recited in Claim 16, wherein scoring said sentence further comprises:

selecting scoring techniques operable for summarizing non-narrative, grammatically incorrect text;

COMPLETE CLAIM SET ACCOMPANYING RESPONSE TO NOTICE OF NON-COMPLIANT
AMENDMENT
Applicant: Maria Castellanos et al.
Serial No.: 09/944,919
Filed: August 31, 2001
Docket No.: H303.239.101
Title: METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT

selecting scoring techniques operable for summarizing narrative, grammatically correct

text; and

using said scoring techniques to score said sentence.

18.   (Currently Amended) The computer system as recited in Claim 17, wherein said method

further comprises generating ~~a~~the summary derived from said scored and ranked sentences.

19.   (Original) The computer system as recited in Claim 11, wherein said method further

comprises selecting a text mining component based upon said data mining operation to be

performed.

20.   (Original) The computer system as recited in Claim 11, wherein said method further

comprises customizing said method by adjusting a parameter value.

21.   (Currently Amended) A computer-useable medium having computer-readable program

code embodied therein for causing a computer system to perform the steps of:

removing an instance of dirty text within said document to produce a cleaned document

having a content; and

performing a data mining operation on said cleaned document to provide a summary of

said content.

22.   (Original) The computer-useable medium of Claim 21, wherein said removing further

comprises replacing an instance of dirty text with a standard term.

23.   (Original) The computer-useable medium recited in Claim 21, wherein said removing

further comprises removing an instance of computer code from said document.

COMPLETE CLAIM SET ACCOMPANYING RESPONSE TO NOTICE OF NON-COMPLIANT
AMENDMENT
Applicant: Maria Castellanos et al.
Serial No.: 09/944,919
Filed: August 31, 2001
Docket No.: H303.239.101
Title: METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT

24.   (Original) The computer-useable medium recited in Claim 21, wherein said removing further comprises removing a table from said document.

25.   (Original) The computer-useable medium recited in Claim 21, wherein said performing a data mining operation further comprises identifying a sentence within said cleaned document by identifying a beginning and an end of said sentence.

26.   (Original) The computer-useable medium recited in Claim 25, wherein said performing a data mining operation further comprises scoring and ranking said sentence.

27.   (Original) The computer-useable medium recited in Claim 26, wherein scoring said sentence further comprises:

selecting scoring techniques operable for summarizing non-narrative, grammatically incorrect text;

selecting scoring techniques operable for summarizing narrative, grammatically correct text; and

using said scoring techniques to score said sentence.

28.   (Original) The computer-useable medium recited in Claim 27, wherein said method further comprises generating a summary derived from said scored and ranked sentences.

29.   (Original) The computer-useable medium as recited in Claim 21, wherein said method further comprises selecting a text mining component based upon said data mining operation to be performed.

30.   (Original) The computer-useable medium as recited in Claim 21, wherein said method further comprises customizing said method by adjusting a parameter value.

COMPLETE CLAIM SET ACCOMPANYING RESPONSE TO NOTICE OF NON-COMPLIANT AMENDMENT
Applicant: Maria Castellanos et al.
Serial No.: 09/944,919
Filed: August 31, 2001
Docket No.: H303.239.101
Title: METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT

31.   (New) A computer-implemented method for mining a document containing dirty text comprising:

producing a cleaned document having a content comprising performing a general cleaning of said document by removing one or more instance of dirty text within said document including instances of misspelling and grammatical errors, and performing a domain and task specific cleaning of said document including removing instances of computer code and tables; and

performing a data mining operation on said cleaned document, including determining a sentence score for each sentence of said cleaned document and ranking the sentences from highest to lowest based on the sentence score;
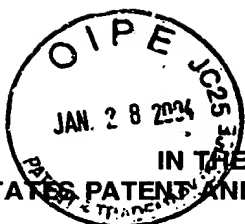
generating a summary of the content of the document using the highest ranked sentences.

32.   (New) The method of claim 31, wherein determining a sentence score for each sentence includes applying a keyword technique to each sentence.

33.   (New) The method of claim 32, wherein determining a sentence score further comprises applying a location technique to each sentence.

34.   (New) The method of claim 32, wherein determining a sentence score further comprises applying a semantic similarity technique to each sentence.

35.   (New) The method of claim 34, wherein the semantic similarity technique comprises:
        generating a vector associated with each sentence; and
        comparing each vector to every other vector, including defining a cosine of an angle between two vectors and using the cosine of the angle between two vectors to determine whether sentences represented by the two vectors are semantically related.

PATENT APPLICATION

ATTORNEY DOCKET NO. ___10007912-1___

2171

UNITED STATES PATENT AND TRADEMARK OFFICE

Inventor(s): Maria Castellanos et al.	Confirmation No.:
Application No.: 09/944,919	Examiner: Wayne P. Amsbury
Filing Date: August 31, 2001	Group Art Unit: 2171
Title: METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT	

RECEIVED

Commissioner for Patents
PO Box 1450
Alexandria, VA 22313-1450

JAN 3 0 2004

Technology Center 2100

TRANSMITTAL LETTER FOR RESPONSE/AMENDMENT

Sir:

Transmitted herewith is/are the following in the above-identified application:

( ) Response/Amendment   ( ) Petition to extend time to respond

( ) New fee as calculated below   ( ) Supplemental Declaration

( ) No additional fee   (Address envelope to "Mail Stop Non-Fee Amendment")

(X) Other: __Response to Non-Compliant Amendment; Claim Set__ _____ (fee $_____)

CLAIMS AS AMENDED BY OTHER THAN A SMALL ENTITY						
(1) FOR	(2) CLAIMS REMAINING AFTER AMENDMENT	(3) NUMBER EXTRA	(4) HIGHEST NUMBER PREVIOUSLY PAID FOR	(5) PRESENT EXTRA	(6) RATE	(7) ADDITIONAL FEES
TOTAL CLAIMS		MINUS		= 0	X $18	$ 0
INDEP. CLAIMS		MINUS		= 0	X $86	$ 0
[ ] FIRST PRESENTATION OF A MULTIPLE DEPENDENT CLAIM					+ $290	$ 0

EXTENSION FEE	1ST MONTH $110.00	2ND MONTH $420.00	3RD MONTH $950.00	4TH MONTH $1480.00	$ 0
				OTHER FEES	$
				TOTAL ADDITIONAL FEE FOR THIS AMENDMENT	$ 0

Charge $___0___ to Deposit Account 08-2025. At any time during the pendency of this application, please charge any fees required or credit any overpayment to Deposit Account 08-2025 pursuant to 37 CFR 1.25. Additionally please charge any fees to Deposit Account 08-2025 under 37 CFR 1.16 through 1.21 inclusive, and any other sections in Title 37 of the Code of Federal Regulations that may regulate fees. A duplicate copy of this sheet is enclosed.

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Commissioner for Patents, Alexandria, VA 22313-1450.

Date of Deposit: **Jan. 26, 2004**

Typed Name: **Steven E. Dicke**

Signature: _Steven E. Dicke_

Respectfully submitted,

Maria Castellanos et al.

By _Steven E. Dicke_

Steven E. Dicke

Attorney/Agent for Applicant(s)
Reg. No. **38,431**

Date: **Jan. 26, 2004**

Telephone No.: **(612) 573-2002**

Rev 10/03 (TransAmd)

- Attach as First Page to Transmitted Papers -